



PvuI (7)
SgfI (6)
MfeI (82)

1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
 101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Bsu36I (291)

Psp1406I (203)
PvuII (239)
Bsu36I (291)

201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCAGGGGCTCGCATCTCTCTTACCGCGCCCGCCGCTACCTGAGGGCC
 301 GCCATCCACGCGGTTGAGTGCCTTCTGCCGCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMI (441)
NaeI (441)
NgoMIV (441)

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGGCTCAACTCTACGCTTTGTTTCGTTT

KasI (535)
AgeI (552)
NcoI (560)
BstEII (555)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACC**GGCGCT**ACCTGAGATCACCGGTCAACGGTCAACGGTGGGGGACAGGCGCTGGATCTCTAGCCTCTGTCCCTT
 601 ACCCAGGCCATGTCCCTTCTGTGCTCCTGTGCTGCTGGTGGTGCCTCGGGGGGCCAGCCCAAGGCTGGCAGGAACACACAGAGCCCCAGGACCT
 13> P R P C P F L L L L L L L V V P R G A Q P Q A G R N H T E P P G P
 701 AACGTTACAGCCACCCCGTGACCCACGATCCCTGTGATCTCTGGAAATGTCAGCACCTCAACAGAAAGTGTCCAGCAGCAGAGACTGAGGGACCCC
 47> N V T A T P V T P T I P V I S G N V S T S T E S A P A A E T E G P

Acc65I (809)
Asp718I (809)
SandI (792)

801 AAAGTGAGAGGTACCCCTCCCTCAGCAGCAGCCCCCTGGGGCCAAAGTGTCCAGGCTGGGCGAGCGTGCAGGTTCCCTTTCCGCTACGGTGG
 80> Q S E R Y P P P S S S S P P G G Q V L T E S G Q P C R F P F R Y G G
 901 CCGCATGCTGACTCCTGTACCTCTGAGGGAAGTGCCTACAGGAAGTGGTGGCTACAACACACAACCTATGACCGAGACCGGGCCTGGGGTACTGTGCA
 113> R M L H S C T S E G S A Y R K W C A T T H N Y D R D R A W G Y C A

BstEII (1002)
NcoI (1092)

1001 GAGGTGACCTGCCTGTGGAAGTCCAGCTATCCTTGACCTTGTGCCTTGGGCCCTGCCTCAATGGGGGCACATGTTCCAGTACACATGACCATGGGT
 147> E V T L P V E G P A I L D P C A S G P C L N G G T C S S T H D H G
 1101 CCTACCACTGCTCTTGCCCTGTCGCTTACAGGCAAGGACTGTGGCAGAGAAATGCTTTGATGAAACACGCTACGAGTATTTTGGAGTGGGGGACCA
 180> S Y H C S C P L A F T G K D C G T E K C F D E T R Y E Y F E V G D H
 1201 CTGGGCCCGTGTGAGCGAGGGACATGTGGAGCAATGTGGCTGTATGGAGGGCCAGGCCGGTGTGAAGACACCCACCACAGCTTGTCTGAGCAGCCCA
 213> W A R V S E G H V E Q C G C M E G Q A R C E D T H H T A C L S S P

PshAI (1341)

1301 TGTCTGAACGGAGGCACCTGCCACCTGATTGTGGGCACAGGGACCGCTGTGCACCTGCCCGTGGGCTATGCTGGGCGTTCTGTAACATTGTTCCCA
 247> C L N G G T C H L I V G T G T S V C T C P L G Y A G R F C N I V P

EcoRI (1482)

1401 CAGAGCACTGCTTCTGGAAATGGTACAGAGTACCGAGGCGTGGCCAGCACCGCTGCCTCGGCCTGAGCTGCCTGGCCTGGAATCTGACCTGCTCTA
 280> T E H C F L G N G T E Y R G V A S T A A S G L S C L A W N S D L L Y

BbrPI (1510)
Eco72I (1510)

1501 CCAGGAGCTGCAGTGGACTCAGTGGCTGCTGCTGCTGCTGCTGCTGGCCTGGGCCCTCACGTTACTGCCGAACCCAGACAAGGATGAGAGGCCTTGGTGC
 313> Q E L H V D S V A A A V L L G L G P H A Y C R N P D K D E R P W C
 1601 TATGTGGTGAAGGACAACGCACTGCTGGGAGTATTGCCGCTGACAGCTGTGAATCCCTGGCCAGAGTCCACTCCCAAAGCCCGAGATCCTAGCAG
 347> Y V V K D N A L S W E Y C R L T A C E S L A R V H S Q S P E I L A
 1701 CCCTGCCCCGAGTCAAGCCCGCTGTGCTCCTACCTGTGGCAAGAGGCACAAGAAGAGGACGTTCTTGAGACCACGCATCATCGGGGCTCATCATCTTT
 380> A L P E S A P A V R P T C G K R H K K R T F L R P R I I G G S S S L

XmnI (1839)
PstI (1886)

1801 GCCTGGCTCACACCCTGGCTGGCTGCCATCTACATTGGGAATAGCTTCTGTGCCGGGAGCCTTGTCCATACCTGCTGGGTAGTGTCTGCAGCCACTGC
 413> P G S H P W L A A I Y I G N S F C A G S L V H T C W V V S A A H C
 1901 TTCGCCAACAGCCCCCAGGGACAGCATCACAGTGGTACTGGTGCAGCACTTCTCAACCGCACCGGATGTGACACAGACATTTGGCATTGAGAAGT
 447> F A N S P P R D S I T V V L G Q H F F N R T T D V T Q T F G I E K

Tth111I (2044)
Eco47III (2078)
AfeI (2078)
BsrBI (2091)

2001 ATGTGCCCTACACCCTGTACTCGGTGTTCAACCCCAACAACCATGACCTTGTCTTGATCCGGCTAAAGAAGAAGGAGAGCGTGTGCTGCTCCGCTCCCA
 480> Y V P Y T L Y S V F N P N N H D L V L I R L K K K G E R C A V R S Q

Bsu36I (2122)

2101 GTTTGTTCACCCATCTGCTGCTGAGGCAGGCTCCTTCCCTACTGGACACAAGTGTGAGATTGACAGGCTGGGGCCACATGGATGAAAATGTGAGC
 513> F V Q P I C L P E A G S S F P T G H K C Q I A G W G H M D E N V S
 2201 AGCTACTCCAACCTCCTGCTGGAGGCACTGGTCCCTCTTGTGCTGACCACAAGTGTAGCAGCCAGAGGTATATGGTGTGACATCAGCCCTAACATGC
 547> S Y S N S L L E A L V P L V A D H K C S S P E V Y G A D I S P N M

NgoMI (2305)
NaeI (2305)
NgoMIV (2305)

2301 TCTGTGCCGCTACTTCCAGTCCGATGCTGCCAGGGGACTCAGGTGGGCCCTTGGTCTGTGAGAAGAATGGTGTGGCTTACCTGTATGGCAT
 580> L C A G Y F D C K S D A C Q G D S G G P L V C E K N G V A Y L Y G I

2401 PvuII (2403) KasI (2423) BstXI (2459)
CATCAGCTGGGGTGTGGCTGTGGGCGCCTCAACAAGCCAGGAGTCTACACCCGTGTGGCCAATTATGTGGACTGGATCAACGACCGTATTCGACCGCCC
613▶ I S W G D G C G R L N K P G V Y T R V A N Y V D W I N D R I R P P
NheI (2554)
2501 AAGCGACCCGTGGCTACGTCTGAACCCCGGTCCCAAGGAAGGATGCTCATAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAA
647▶ K R P V A T S •

2601 HpaI (2692)
CCACAAC TAGAATGCAGTGA AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAA

MfeI (2703) EcoRI (2788)
2701 CAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAAATCTAAAAT
2801 ACAGCATAGCAAAACTTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCA
2901 TTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGTATTTTCCCAAGTTTGAAC TAGCTCTTCATTTCTTTATGTTTTAAATGCAC

SspI (3027) SmaI (3041)
3001 TGACCTCCACATTCCTTTTTAGTAAAATATTAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAATCCAGATGCT
3101 CAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTA
3201 GTTCTCGGTGACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAG
140▶ N R T Y K L P I L E E I T T K V L K G N M E I L V F C D P A Y D S

SacI (3302) BstXI (3331)
3301 ATGAGCTCTGCACATGCCACAGGGGTGACCACTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCCCT
106▶ I L E R C M G C P S V V R I S R D V E D S Y P H R V A V I T D F D K
3401 TCTGCCGTTGCTCAGCAGACCAATGGCAATGGCTTCAGCAGACAGCAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCC
73▶ Q G N S V A S G I A I A E A C V T V R G I Y A E I H V A S I I E G
3501 AGTCTGGTCTGATGGCCGCCGACATGGTGTCTGTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGA
40▶ T K T R I A A G V H H K N D E Y L M T I K E T A V E V L E L D Q Q

BspHI (3616) AseI (3674)
XmnI (3608) VspI (3674)
3601 GAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCGGATATACTATGCGGATGATTAATTGTCAAACAGCGTGGATGG
64▶ S I N F T K M

SacI (3731)
3701 CGTCTCCAGC TATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGCGGAGTTG

SpeI (3829)
3800 TTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGC

SnaBI (3957)
Eco105I (3957)
3899 TATCCAGCCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCA

NdeI (4062)
3999 TGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCAGT
4099 TTACCGTAAATACTCCACCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTG

PstI (4241)
SdaI (4240) PacI (4248)
4199 GCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCC T G C A G G T T A A T T A A G A A C A T G T G A G C A A A A G G C C A G A A A A G G C C A G G A A C C G T
4297 A A A A G G C C G G T T G C T G G C G T T T T C C A T A G G C T C G C C C C C T G A C G A G C A T C A C A A A A T C G A C G C T C A A G T C A G A G G T G G C G A A A C C C G A C A G G A C
4397 T A T A A G A T A C C A G G C G T T T C C C C T G G A A G T C C C T C G T G C G C T C C C T G T T C C G A C C T G C C G T T A C C G G A T A C T G T C C G C T T T C C C T T C G G G

ApaLI (4572)
4497 AAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCC
4597 GACCCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCCAAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG
4697 CGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT
4797 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAA

PacI (4988)
4897 AAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTACGTTAAGGGATTTTGGTCATGGCTAGTTAATTA

EagI (5008)
SwaI (4997) NotI (5007)
4997 CATTTAAATC AGCGGCGCAATAAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAAC
5097 AAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAAGTCAAGTGCAAGTGCCAGAACATTTCTCTATCGAA